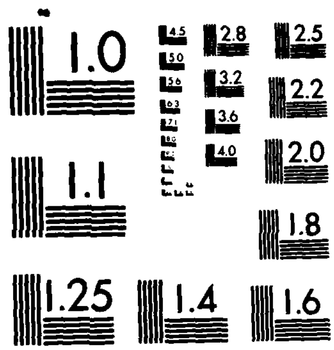
AD-A163 049 DECISION PROCEDURES(U) STANFORD UNIV CA DEPT OF COMPUTER SCIENCE M L GINSBERG MAY 85 STAN-CS-85-1064 N00014-81-K-0004 1/1

UNCLASSIFIED F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

May 1985

Report No. STAN-CS-85-1064
Also numbered KSL-85-25

Decision Procedures

by

Matthew L. Ginsberg

Contract No. N00014-81-K-0004

Department of Computer Science

Stanford University
Stanford, CA 94305

Decision Procedures

Matthew L. Ginsberg

THE LOGIC GROUP
KNOWLEDGE SYSTEMS LABORATORY
Department of Computer Science
Stanford University
Stanford, California 94305

Decision Procedures

Abstract

Distributed artificial intelligence is the study of how a group of individual intelligent agents can combine to solve a difficult global problem; the usual approach is to split the original problem into simpler ones and to attack each of these independently. This paper discusses in very general terms the problems which arise if the subproblems are *not* independent, but instead interrelate in some way. We are led to a single assumption, which we call *common rationality,* that is provably optimal (in a formal sense) and which enables us to characterize precisely the communication needs of the participants in multi-agent interactions. An example of a distributed computation using these ideas is presented.

§1. Introduction

The thrust of research in distributed artificial intelligence (DAI) is the investigation of the possibility of solving a difficult problem by presenting each of a variety of machines with simpler parts of it.

The approach that has been taken has been to consider the problem of dividing the original problem: what subtasks should be pursued at any given time? To which available machine should a given subtask be assigned? The question of how the individual machines should go about solving their subproblems has been left to the non-distributed AI community (or perhaps to a recursive application of DAI techniques).

The assumption underlying this approach—that each of the agents involved in the solution of the subproblems can proceed independently of the others—has recently been called into question [2,3,6,7,10]. It has been realized that, in a world of limited resources, it is inappropriate to dedicate a substantial fraction of those resources to each processor. The increasing attractiveness of parallel architectures in which processors share memory is an example of this: memory is a scarce resource.

Automated factories must inevitably encounter similar difficulties. Are the robots working in such factories to be given distinct bins of component parts, and non-overlapping regions in which to work or to travel from one area of the factory to another? It seems unlikely.

My intention in this paper is to discuss these issues at a very basic (i.e., formal) level. I will be interested in situations where:

(1) A global goal has been replaced by a variety of local ones, each pursued by an individual agent or process, and
(2) The actions of the individual agents may interact, in that success or failure for one such agent may be partially or wholly contingent upon an action taken by another.

The first of these is a massive disclaimer. I will *not* be concerned with the usual problem of dividing the global goal, but will assume that this has already been done. My intention is merely to remove the constraint that the subproblems or subagents cannot interact; the problem of subdividing problems in the absence of this constraint is outside the intended scope of this paper.

The second remark above explicitly allows for "partial" success or failure—corresponding, for example, to the speed with which a given subgoal is achieved.

In the case where the agents do not interact, we will assume that each knows enough to evaluate the results of its possible actions. For agent i, this evaluation will be incorporated in a payoff function p which assigns to any alternative m the value of that course of action. Thus if M represents the set of all of i's possible courses of action, p is simply a function

$$p : M \rightarrow \mathbb{R}. \tag{1}$$

That the range of p is $\mathbb{R}$ as opposed to $\{0, 1\}$ reflects our allowing for a varying degree of success or failure.

Determining the function p is a task that lies squarely within the province of non-distributed AI research. Given such a function, selecting the alternative m which maximizes $p(m)$ is straightforward.

In the absence of interaction, the problems of the individual agents would now be solved. If the success or failure of some agent depends on actions taken by another, however, this is not the case: the function p in (1) will have as its domain the set of actions available to *all* of the agents, as opposed to a single one.

This sort of problem has been discussed extensively by game theorists. They do, however, generally make the assumption that the agents have common knowledge of each others' payoff functions. This need not be valid in an AI setting.

We will address this problem in due course; our basic view is that the fundamental role of communication between interacting intelligent agents is to establish an agreed payoff function as described in the last paragraph. Before turning to this, let us examine in an informal way some situations in which interaction is important.

The first is one which we will refer to as *coordination*. Suppose that two robots, one in Boston and the other in Palo Alto, decide to meet in Pittsburgh to build a widget. Now, building a widget is a complicated procedure unless you happen to have both a zazzyfrax and a borogove, although either in isolation is useless. Zazzyfraxen are available only in New York, and borogoves only in San Francisco; should the robots stop on their travels to acquire their respective components?

The answer is clearly that they should; note, however, that for either robot to do so involves a great many assumptions about the other robot. The Palo Alto robot needs not only to know about the availability of zazzyfraxen in New York—he must also assume that the Boston robot knows about the borogoves in San Francisco, *and* that the Boston robot knows the Palo Alto robot knows about the zazzyfraxen, and so on. Halpern and Moses [8] have described this sort of situation as *common knowledge*.

As we remarked earlier, common knowledge of this sort is presumed by the game theorists in their assumption of common knowledge of the payoff function. But even this is not enough to ensure coordination of the robots' actions: the Palo Alto robot must assume that the Boston robot is sensible enough to stop in New York, which implies that the Boston robot knows *he* is sensible enough to stop in San Francisco, and so on. So even common knowledge of the payoff function is not enough: some sort of common knowledge of the methods by which the robots select their actions is also required.

The game theorists deal with this last requirement by fiat, assuming [13] that any "rational" agents will coordinate their actions in such a situation. This seems unnecessarily

ad hoc, and also appears to conflict with the philosophy that each agent strives to achieve a purely local goal.

Our robots somehow overcome these difficulties, starting a successful widget business in Pittsburgh. At some point, however, their zazzyfrax breaks, and they arrange with the New York factory to exchange five widgets for a replacement. Since the robots need the zazzyfrax as quickly as possible and the New York factory is equally anxious to get its widgets, it is decided that each group will ship their contribution to the exchange immediately, trusting the other to act similarly.

In the absence of legal sanctions, why should either agent ship the goods? The result of this, however, is for there to be no transaction—an outcome worse for *both* agents than if they had acted as agreed.

The problem is one of *cooperation*; the scenario we have presented is Hofstadter's [9] reformulation of the well-known *prisoner's dilemma*.

The game theorists have no answer in this case. Not only does the pursuit of a local goal offer no incentive for cooperation, it actively *discourages* it. The robots' business, having become too dependent on the zazzyfrax technology, goes bankrupt.

§2. Solutions to problems of interaction

Our example may have been far-fetched, but the problems are real. The coordination problem has straightforward analogs in automated factories, as does the cooperation one: why should one robot or agent stop to arrange delivery of a part required by another when doing so does not further its own local ambitions?

A variety of solutions has been proposed. The one which has received the most attention in the AI literature [2,3,12] has been to ensure that the agent responsible for dividing the original problem into more manageable smaller ones does so in a way which ensures that the subtasks do not interact. We have already remarked that this may be overly constraining in many situations.

It has also been suggested that the agents communicate their intentions to one another. Rosenschein [15] has demonstrated, however, that there is no reason for independent agents to communicate honestly. Since a statement of the form, "I intend to do x," is virtually unprovable, the value of such communication is extremely limited. Indeed, it was the non-binding nature of communication that led to the downfall of the entrepreneurial robots of the last section. Rosenschein and Genesereth [14] have suggested remedying this by making the commitments binding, but it is not clear how this can be enforced.

The game theorists use different approaches. We have already remarked upon their adoption of a definition of rationality that considers coordinated action; it unfortunately does not seem possible to formalize this in a way that is clearly consistent with the "local utility maximization" that is generally assumed. The game theorists themselves have remarked upon this [11].

Cooperation is harder to come by than is coordination. The game theorists achieve it by assuming there is some facility for retribution; this amounts to assuming that the given interaction is only one of a series. Axelrod [1] has shown that cooperation can evolve under these assumptions.

Again, the view that an agent's actions should be governed by a concern for its future welfare is not in keeping with the idea that such an agent should act to maximize a purely

local payoff. We have similar objections to schemes that include an *ad hoc* "altruism factor" in the payoff function.

§3. Notation

In order to formalize a solution to these difficulties, we need first to formalize the problem. We begin by generalizing (1). There, for each agent i, we had a function

$$p_i : M_i \to \mathbb{R},$$

where M_i is the set of allowable "moves" for player i. In the interactive case, this generalizes to a function

$$p_i : \prod_{i \in P} M_i \to \mathbb{R}, \tag{2}$$

where P is the set of all players involved in the interaction.

For $S \subset P$, we will denote $P - S$ by $\bar{S}$; we will also write simply i instead of $\{i\}$ where no confusion is possible. (Thus $\bar{\imath} = P - \{i\}$, for example.) We also write M_S for $\prod_{i \in S} M_i$, so that (2) becomes

$$p_i : M_P \to \mathbb{R}.$$

The various p_i can be collected into a single *payoff function*

$$p : P \times M_P \to \mathbb{R}. \tag{3}$$

We will identify a game or interaction g with the associated payoff function as in (3).

We will denote by m_S an element of M_S; this is a collective move for the players in S. To $m_S \in M_S$ and $m_{\bar{S}} \in M_{\bar{S}}$ correspond an element $\vec{m}$ of M_P. The payoff function in (3) can now be interpreted by noting that for $(i, \vec{m}) \in P \times M_P$, $p(i, \vec{m})$ is the payoff to player i if the joint move $\vec{m}$ is made. We can now see that a game is *non-interactive* if and only if it satisfies:

$$m_i = n_i \quad \Rightarrow \quad p(i, \vec{m}) = p(i, \vec{n});$$

in other words, the payoff to each player depends only on that player's action.

In cases where only two agents are interacting, we will use a matrix notation to represent the payoff function. Here is an example:

<table>
<tr><th></th><th>C</th><th>D</th></tr>
<tr><td>A</td><td>1
3</td><td>2</td></tr>
<tr><td>B</td><td>5
2</td><td>1
0</td></tr>
</table>

The first player selects one of the two rows by choosing A or B; the second one of the columns by choosing C or D. The payoffs correspond to the numbers in the associated boxes: if move AC is selected, for example, the row player above would receive a payoff of 3 while the column player would receive 1. The single entry "2" in the AD position corresponds to an identical payoff of 2 for both players.

Here is the payoff function for a non-interactive game:

	C	D
A	3, 5	3, 1
B	1, 5	1

Here is one requiring coordination:

	C	D
A	7	4
B	5	6

Games such as this, where one outcome (AC above) is preferred by all players to any other, are referred to as "no-conflict games" in the game theory literature [13].

Finally, here is the prisoner's dilemma. Each agent can "cooperate" (C) or "defect" (D):

	C	D
C	3	0, 5
D	5, 0	1

Independent of the other's action, each agent is better off defecting than cooperating. Unfortunately, the payoff resulting from mutual defection is worse for *both* agents than that corresponding to mutual cooperation.

§4. Rationality

Game theory literature generally examines single games and considers what it means for specific moves in these games to be rational. "Rational" here is generally only defined informally, having to do with maximizing the player's return under some ill-defined assumptions.

It seems to me that, rather than the particular move being rational (or not), it is the *analysis leading to the move* which should be examined. So I will try to develop a framework that lets me discuss these analyses rather than simply the moves they select.

Let P be a fixed group of players. Given a game g, let us denote by g_i the moves for i which are legal in g. We also denote the set of all games by G, and take G_i to be $\cup_{g \in G} g_i$, so that G_i is the collection of all moves for i which are legal in *some* game. We will assume throughout that any game obtained by permuting the players in some particular game $g \in G$ is also in G, so that $G_i = G_j$ for all i and j.

We now define a *decision procedure* for player i to be a function

$$D_i : G \to G_i$$

such that $D_i(g) \in g_i$ for all $g \in G$. In other words, D_i assigns to each game a specific legal move for i in that game and therefore encodes the process whereby i selects his move. For a fixed group S of players, we denote $\prod_{i \in S} D_i$ by D_S. Thus

$$D_S(g) = \prod_{i \in S} D_i(g);$$

D_S is the "collective" decision procedure used by the players in S to select a joint move.

Suppose now that player i uses his decision procedure D_i to select a move m_i in a game g, and that p is the payoff function associated to this game. The other players in the game use their collective decision procedure $D_{\bar{i}}$ to select a move $m_{\bar{i}}$ and the resulting payoff to i is therefore

$$pay(i, D_i, g) = p(i, \vec{m}). \tag{4}$$

The function *pay* here gives the payoff to i in the game g if he uses the decision procedure D_i; note that the appearance of $\vec{m}$ in (4) means that *pay* is implicitly a function of all of the D_j's, and not just D_i.

If we knew the expected distribution of games, we could amalgamate (4) over all games to get the expected payoff corresponding to the decision procedure D_i generally. Although this is an extremely strong assumption, we will have use of it, so here goes: We need a *density function* ρ_i which assigns to each game $g \in G$ its likelihood of occurring; $\sum_{g \in G} \rho_i(g) = 1$. The expected payoff corresponding to D_i is now given by

$$pay(i, D_i) = \sum_{g \in G} \rho_i(g) pay(i, D_i, g). \tag{5}$$

Again, this has implicit dependence on the decision procedures of the other players as well.

Under these strong assumptions, it is clear that rationality corresponds to maximizing (5). The idea we want to capture is that a decision procedure is *irrational* if it results in a payoff that is provably suboptimal—recall, though, that the implicit dependence of (5) on $D_{\bar{i}}$ leads to a potential ambiguity in a definition such as that D_i is irrational if there exists another decision procedure C_i such that

$$pay(i, D_i) < pay(i, C_i).$$

We will resolve this ambiguity by quantifying over the $C_{\bar{i}}$ and D_i which appear implicitly in the above equation, but in doing so, want to allow for the possibility of encoding some restrictions on the decision procedures of the other players. Retaining as much generality as possible, we will assume the existence of some predicate *allowed*$(C_{\bar{i}}, D_{\bar{i}})$, and now say that a decision procedure D_i is *globally irrational* (with respect to *allowed*) if there exists a decision procedure C_i such that, *for all* $C_{\bar{i}}$ *and* $D_{\bar{i}}$ *with allowed*$(C_{\bar{i}}, D_{\bar{i}})$,

$$pay(i, D_i) < pay(i, C_i). \tag{6}$$

In general, of course, the density function ρ_i will not be known; we therefore need a notion of rationality depending only upon the local payoffs appearing in (4). To this end, we define a decision procedure D_i to be *locally irrational* (or simply *irrational*), again with respect to *allowed*, if there exists a decision procedure C_i such that, for all $C_{\bar{i}}$ and $D_{\bar{i}}$ with *allowed*$(C_{\bar{i}}, D_{\bar{i}})$,

$$pay(i, D_i, g) \leq pay(i, C_i, g) \tag{7}$$

for all games g, with the inequality being strict in at least one case. In other words, a decision procedure is irrational if there is another decision procedure which is better in some specific game, and no worse in all others. We will call a decision procedure *rational* (or *locally rational*) if it is not irrational, and *globally rational* if it is not globally irrational.

Note that we do *not* call a decision procedure irrational merely because it does not achieve an optimal payoff in *every* game: it is only if a better payoff could be achieved in some game *without reducing the payoffs in other games* that the definition applies. Indeed, we shall see in section 8 that there are situations where it is in fact necessary to accept a suboptimal payoff in certain games.

Lemma 4.1. *Any globally rational decision procedure is locally rational.*

Proof. For any allowed decision procedures $C_{\bar{i}}$ and $D_{\bar{i}}$, summing (7) over all games reproduces (6). □

The results of the next section deal with conditions which must be satisfied by any locally rational decision procedure. The point of the lemma is that a globally rational decision procedure must meet these requirements as well.

The nature of the definitions of rationality depends, of course, on the *allowed* predicate. If, for example, we knew the decision procedures of the other players, we could restrict *allowed* to these decision procedures, so that our decision procedure could be easily determined by the definition of local rationality (7). Essentially, we would be moving with complete knowledge of the other players' choices of action.

In practice, of course, this will not be the case, and *allowed* will hold for a variety of decision procedures other than the ones the other players are actually using. This will lead to some freedom in our own (presumed rational) choice of decision procedure; our interest will be in determining the nature and amount of this freedom given a variety of (incomplete) assumptions regarding the decision procedures of the other players or agents.

At this point, we have made no such assumptions at all. The fact that *allowed* is a function of *two* decision procedures, for example, allows us to consider situations where the decision procedures of the other agents are not independent of our own. In fact, the ability to drop this independence assumption is the pragmatic reason we are working not with the moves themselves, but with the analyses which lead to them. By working in this fashion, it is possible to assume that the agents have knowledge of each others' strategies without being led to the circular arguments that otherwise pervade this sort of analysis.

As an example, consider the following well-known "paradox": an alien approaches you with two envelopes, one marked "\$" and the other "¢". The first envelope contains some number of dollars, and the other the same number of cents. The alien is prepared to give you the contents of either envelope.

The catch is that the alien, who is omniscient, is aware of the choice you will make. In an attempt to discourage greed on your part, he has decided to put one unit of currency in the envelopes if you will pick "\$", but one thousand units if you will pick "¢". Bearing in mind that the alien has decided upon the contents of the envelopes *before* you pick one, which should you select? (If this example is insufficiently compelling for an AI audience, multiply all of the figures by a million.)

Here are the payoffs for this game; the alien's payoffs simply reflect his desire to teach you the desired lesson. (He can make all the money he needs in the stock market, anyway.)

	1000	1
\$	0 1000.	1 1.
¢	1 10.	0 0.01

Since the payoff for \$ is greater than that for ¢ for either of the alien's options, any of the conventional game-theoretic analyses will lead us to select the dollar envelope, and we will presumably receive only \$1 as a result.

The decision procedure paradigm is flexible enough to handle this situation. If we are player 1 and the alien is player 2, we have that the only allowed D_2 is:

$$D_2(g) = \begin{cases} 1000, & \text{if } D_1(g) = ¢, \text{ and} \\ 1, & \text{if } D_1(g) = \$. \end{cases} \tag{8}$$

Inserting this into (4) gives $pay(1, ¢, g) = 10$ and $pay(1, \$, g) = 1$, allowing us to conclude from (7) that it is indeed irrational to pick \$ in this game.

In no way have we avoided the paradox of the alien's omniscience; we have merely found a way to *describe* this omniscience in the decision procedure (8). We will see in the next section that the "case analysis" argument which would have led us to select \$ in the above game is valid if the decision procedures of the various players are independent.

It is a small step from the scenario we have described to one with a more serious difficulty. Suppose that the player 1 in the above game, instead of being you or I, is another omniscient alien. The new alien, in an attempt to encourage the original one to be more cautious with its money, decides to select \$ if the offering alien puts one thousand units of currency in the envelopes, and to select ¢ otherwise. What happens?

Here are the payoffs for the new game:

	1000	1
\$	0 1	1 0
¢	1 0	0 1

The situation appears to be circular; in fact the descriptions of the two aliens are inconsistent. The decision procedure of the new alien is supposedly given by:

$$D_1(g) = \begin{cases} ¢, & \text{if } D_2(g) = 1, \text{ and} \\ \$, & \text{if } D_2(g) = 1000. \end{cases}$$

This in in clear conflict with (8), in that there are no allowed $C_{\bar{\imath}}$ or $D_{\bar{\imath}}$ for this game. (Is this example an argument for monotheism?)

We will be interested in a variety of assumptions regarding the decision procedures of the other agents; they will fall into two distinct types. The first consists of *rationality* assumptions: we can assume that the other agents are locally or globally rational. Alternatively, we can make *behavior* assumptions, assuming that the other agents' behavior is independent of our own, or that it is the same as our own.

Let us deal with the rationality assumptions first. We will say that a joint decision procedure D_S is rational if it is the product of individual decision procedures which are all rational, and now can define:

1a. **Separate (local) rationality:** $C_{\bar{\imath}}$ not locally rational $\Rightarrow \neg allowed(C_{\bar{\imath}}, D_{\bar{\imath}})$, and similarly for $D_{\bar{\imath}}$.

1b. **Separate global rationality:** $C_{\bar{\imath}}$ not globally rational $\Rightarrow \neg allowed(C_{\bar{\imath}}, D_{\bar{\imath}})$, and similarly for $D_{\bar{\imath}}$.

The appearance of *allowed* in the rationality definitions may appear to make these last definitions circular, but this is not the case. The reason for this is that as the *allowed* set gets smaller, more and more decision procedures will be defined to be irrational by (6) and (7). The separate rationality assumptions can therefore be used to produce increasingly sharp limits on the decision procedures of the agents involved. This can be formalized by defining the set of rational decision procedures for player i to be the maximal fixed point of an operator defined in terms of (6) or (7). (The example following corollary 5.2 is an example of this.)

The behavior assumptions are more difficult to capture formally. When we say, "the other agents will behave as I do," for example, we mean that the other agents would act similary *were they to find themselves in my current situation.*

This notion depends on that of "permuting" the players in a game in order to examine that game from another player's point of view. In order to capture this idea, suppose that there are n players in P, and denote by S_n the group of permutations on an n-element set. Given a game with payoff function p, and a fixed permutation $\sigma \in S_n$, we can define a permuted game p_σ with payoff function given by

$$p_\sigma(\sigma(i), \sigma(\vec{m})) = p(i, \vec{m}).$$

Since we are identifying games and payoff functions, it follows from this that a permutation σ induces a mapping (also to be denoted σ) from games to games.

To get a feel for this, we might say that a decision procedure C_i is *unbiased* if $C_i(g) = C_i(\sigma(g))$ whenever $\sigma(i) = i$. In other words, a shuffling of the other players in the game does not affect i's choice of action; he is uniform in his treatment of the other agents.

Slightly more generally, we might fix $S \subset P$ and assume that $\sigma C_S(g) = C_S(\sigma(g))$ whenever $\sigma(S) = S$. The implication of this is not only that the group S is unbiased toward the remaining players in $\bar{S}$, but also that the behaviors of the individuals within the group are identical in the above sense.

We now define:

2a. **Common behavior:** If, for some σ, $\sigma C_P(g) \neq C_P(\sigma(g))$, then $\neg allowed(C_{\bar{\imath}}, D_{\bar{\imath}})$, and similarly for $D_{\bar{\imath}}$.

This assumption (which cannot be described in the conventional game-theoretic formalism) is peculiarly appropriate to an AI setting; it would be straightforward to equip potentially interacting agents with matching decision procedures. We will see in section 6 that it is in fact sufficient to equip the agents with only the common behavior *assumption* (theorem 6.2), and also that this is, at least in some sense, a good idea (theorem 6.3).

We also have:

2b. **Independent behavior:** $C_{\bar{i}} \neq D_{\bar{i}} \Rightarrow \neg allowed(C_{\bar{i}}, D_{\bar{i}})$. This amounts to assuming that each agent's decision procedure does not influence the others. This assumption is generally made in the game theory literature, and is equivalent to assuming that the moves of the other players are fixed in advance (since their decision procedures are), although the exact nature of these moves remains unknown.

Separate rationality and independent behavior are independent but consistent; there are decision procedures which are rational under one assumption but irrational under the other. We will refer to the combination of separate rationality and independent behavior as *individual rationality* (or perhaps as *individual global rationality*).

Lemma 4.2. *Common behavior implies separate rationality if at least one of the players has a rational decision procedure.*

Proof. If some decision procedure D_i leads to a payoff $pay(i, D_i, g) = p(i, \vec{m})$ for player i under the common behavior assumption, the payoff to some other player $j = \sigma(i)$ in the permuted game $\sigma(g)$ will be

$$\begin{aligned} pay(j, D_j, \sigma(g)) &= p_\sigma(\sigma(i), D_j(\sigma(g))) \\ &= p_\sigma(\sigma(i), \sigma(D_i(g))) \\ &= p(i, D_i(g)) = pay(i, D_i, g). \end{aligned}$$

It follows that if a decision procedure C_i produces a universal improvement for i, the permutation $\sigma C_i \sigma^{-1}$ will produce a universal improvement for j. Thus, if D_i were irrational for i, $\sigma D_i \sigma^{-1}$ would be irrational for j. The conclusion follows. □

Note that it seems impossible to strengthen lemma 4.2. For example, it is not necessarily the case that a decision procedure that is rational for one agent is rational for all others. The presence of a third player whose actions are known to be biased in favor of one of the others can introduce asymmetry of this sort—we approach problems differently if there are additional agents present upon whose support and cooperation we can rely.

It is also not the case that common behavior implies separate *global* rationality. Differing density functions can easily result in a decision procedure that leads to a global improvement for one agent while being a disaster for another. Suppose, however, that we say that the various density functions ρ_i *match* if $\rho_{\sigma(i)}\sigma = \rho_i$ for all $\sigma \in S_n$. We then have the following:

Lemma 4.3. *Common behavior implies separate global rationality if at least one of the players has a globally rational decision procedure and all of the players have matching density functions.*

Proof. If the density functions match, the global payoffs to all of the players are identical; the conclusion follows. (The proof that the global payoffs are identical is embedded in the proof of theorem 6.3.) □

We will refer to the combination of common behavior and separate rationality as *common rationality* (or perhaps as *common global rationality*).

A few more words about matching density functions are probably in order before we return to local issues. If we are interested in situations where all of the agents *agree* about the likelihood of various interactions, then the density functions will match if and only if $\rho = \rho\sigma$ for all σ, so that the likelihood of a game is not changed by permuting it. In other words, equal density functions will match just in case each agent is as likely to find himself in any particular situation as is any other. The more general formulation amounts the statement that the agents *believe* themselves to be equally likely to find themselves in specific interactions—for example, each might believe that interactions unfavorable to him were more likely than others. These differing density functions would still match if the amount of pessimism were uniform among the various agents.

§5. Local rationality

In this section we investigate the consequences of our definitions of rationality. We will assume throughout that g is some fixed game with payoff function p.

We define a single move m_i to be *rational* in a game g if there is a rational decision procedure D_i with $D_i(g) = m_i$, and now have:

Theorem 5.1 (Case analysis). *Assuming independent behavior, if for some moves c_i and d_i, for all $m_{\bar{\imath}}$,*

$$p(i, d_i \times m_{\bar{\imath}}) < p(i, c_i \times m_{\bar{\imath}}),$$

then d_i is not rational in g.

Proof. Suppose that $D_i(g) = d_i$, and let C_i be the decision procedure given by

$$C_i(g') = \begin{cases} D_i(g'), & \text{if } g' \neq g; \\ c_i, & \text{if } g' = g. \end{cases}$$

For any allowed $C_{\bar{\imath}}$ and $D_{\bar{\imath}}$, $C_{\bar{\imath}} = D_{\bar{\imath}}$; it follows that if we set $m_{\bar{\imath}} = C_{\bar{\imath}}(g) = D_{\bar{\imath}}(g)$,

$$pay(i, D_i, g) = p(i, d_i \times m_{\bar{\imath}}) < p(i, c_i \times m_{\bar{\imath}}) = pay(i, C_i, g),$$

while $pay(i, D_i, g') = pay(i, C_i, g')$ for $g' \neq g$, so that the decision procedure D_i is irrational. □

It is an invalid application of this theorem that led us to choose \$ in the paradox of the last section. Of course, the formulation of the interaction there is such as to indicate clearly that the independent behavior assumption does *not* hold.

The dependence of the proof on the independent behavior assumption lies in part in the statement that $pay(i, D_i, g') = pay(i, C_i, g')$ for $g' \neq g$. In order to conclude this, we needed to know that the other agents would not change their actions in g' in response to our selecting the move c_i in g.

Corollary 5.2 (Iterated case analysis). *Assuming individual rationality, if for some moves c_i and d_i, for all rational $m_{\bar{i}}$,*

$$p(i, d_i \times m_{\bar{i}}) < p(i, c_i \times m_{\bar{i}}),$$

then d_i is not rational in g.

Proof. Clear. □

If this result allows us to conclude that there is a unique rational move in some game g, this is known as the *solution in the complete weak sense* in the game theory literature.

The consequences of case analysis are quite well known. Consider the example used to introduce our payoff matrix notation:

	C	D
A	1 3	2
B	5 2	1 0

Independent of the column player's choice, the row player will be better off if he makes move A, since his payoff will be 3 as opposed to 2 if the other player selects C, and 2 as opposed to 0 if D is chosen. Case analysis therefore implies that A is the only rational move for the row player in this game.

There is no such implication for the column player. Assuming individual rationality, however, he will realize that the row player can be counted on to make move A, and will therefore respond with D (receiving a payoff of 2 instead of 1).

Note that the column player would have benefitted, at no cost to the row player, if BC had been selected instead of AD. The prisoner's dilemma is an extension of this idea:

	C	D
C	3	5 0
D	0 5	1

Case analysis leads each player to defect, leading to a result which is unfortunate for both of them.

The difficulty is a consequence of the fact that the agents are pursuing their goals too blindly: they are employing the independent behavior assumption in a situation where it is inadvisable to do so. In general, the consequences of theorem 5.1 will make the independent behavior assumption inconsistent with any cooperative or altruistic behavior.

Common behavior provides an answer:

Theorem 5.3 (Cooperation). *Assume common behavior and local rationality, and suppose that g has moves $\vec{c}$ and $\vec{d}$ such that*

$$p(j, \vec{d}) \leq p(j, \vec{c})$$

for all players j, with the inequality being strict in at least one case. Then if P denotes the set of all agents in the interaction, and $\vec{c}$ is possible as the outcome of a common decision procedurefor the agents in P, $D_P(g) \neq \vec{d}$. In other words, the collective move $\vec{d}$ is avoided.

Proof. Suppose that $D_P(g) = \vec{d}$, so that $D_i(g) = d_i$ for all players i. Now let G' be the set of all games obtained by permuting the players in P, and consider the decision procedure C_i obtained by modifying D_i so that the move $\vec{c}$ and its permuted versions are obtained for the games in G'. The hypotheses of the theorem will allow us to apply (7) to conclude that D_i is irrational. □

In other words, if mutual defection is to everyone's disadvantage, then at least one player will cooperate. For an interaction such as the prisoner's dilemma which is symmetric (in that $p_\sigma = p$ for all σ), it follows that all of the players cooperate. This conclusion has an analog in the informal arguments of [4] and [9].

No-conflict games are handled as a special case of this result:

Corollary 5.4 (Coordination). *Assume common behavior, and suppose that g has a move $\vec{c}$ such that for any $\vec{d} \neq \vec{c}$,*

$$p(j, \vec{c}) \geq p(j, \vec{d})$$

for all players j, with the inequality being strict for at least one j. Then $D_P(g) = \vec{c}$.

Proof. Apply theorem 5.3 to eliminate each of the alternatives. □

Returning to 5.3 itself, note that the theorem does *not* imply that all of the players cooperate; the possibility of one agent cooperating while the others defect (undoubtedly beneficial to the defectors and disastrous for the cooperator) is typical of the "altruism" allowed under the common behavior assumption. Consider the following non-result:

Non-theorem 5.5 (Restricted case analysis). *Assume minimal rationality, and suppose that there exist c_i and d_i such that, for all $c_{\bar{i}}$ and $d_{\bar{i}}$,*

$$p(i, \vec{d}) < p(i, \vec{c}).$$

Then d_i is irrational in g.

Proof? If $D_i(g) = d_i$, consider the decision procedure given by

$$C_i(g') = \begin{cases} D_i(g'), & \text{if } g' \neq g; \\ c_i, & \text{otherwise.} \end{cases}$$

This leads to $pay(i, C_i, g) > pay(i, D_i, g)$ independent of the nature of the *allowed* predicate. □

The problem is that we cannot show that

$$pay(i, C_i, g') \geq pay(i, D_i, g') \tag{9}$$

for $g' \neq g$. Independent behavior allows us to conclude this, of course; witness theorem 5.1. But under the common behavior assumption, if the "better" move c_i forces a reduced payoff for some other player, then (9) will specifically *not* hold for some permutation of g. This of course does not mean that C_i is irrational; it merely means that C_i cannot be used to prove that D_i is.

Restricted case analysis is in fact independent of the assumption of common behavior, although it is consistent with it.

§6. Global rationality

In this section we will investigate the interplay between the common behavior assumption and the definition of global rationality (6). Before doing so, however, consider the following symmetric game:

	A	B
A	0	1
B	1	0

Common behavior does very poorly here, resulting in a payoff of 0 for both players. The difficulty lies in the symmetry—if the payoffs are replaced by $1+\epsilon$ and $1-\epsilon$, this difficulty will be avoided. We therefore define a game g with payoff function p to be *locally disambiguated* (or simply *disambiguated*) if, $p_\sigma \neq p$ for every non-trivial permutation σ of the players in g.

Recall that we apply the permutation σ to both the player *and* the move. Thus the following (symmetric) game is not locally disambiguated:

	A	B
A	0	$_{1}\;^{2}$
B	$_{2}\;^{1}$	0

while this game *is* locally disambiguated:

	A	B
C	0	$_{1}\;^{2}$
D	$_{1}\;^{2}$	0

To see that this last game is not symmetric, all we need do is note that the column player has a fundamental advantage in it.

Note also that, in the presence of the common behavior assumption $D_j = \sigma D_i \sigma^{-1}$, the payoff function (4) has no unexpressed dependence on the decision procedures of the other players, since there is already an explicit dependence on D_i in $pay(i, D_i, g)$.

For a fixed game g, we define the *orbit* of g, denoted $o(g)$, to be the set of all games obtained by permuting the players in g. We also extend the payoff function to orbits, defining

$$\overline{pay}(i, D_i, g) = \sum_{g' \in o(g)} \rho_i(g') pay(i, D_i, g'). \tag{10}$$

A game g will be called *globally disambiguated* if $\overline{pay}(i, D_i, g) \neq \overline{pay}(i, C_i, g)$ whenever $D_i(g) \neq C_i(g)$. Note that, as a result of the appearance of ρ_i in (10), this definition is dependent upon the values of the density function.

Theorem 6.1. *Assume common behavior. Then there is a unique globally rational move in any globally disambiguated game.*

Proof. We can assume without loss of generality that G, the set of all games, is equal to $o(g)$ for some single game g. Now $pay(i, D_i) = \overline{pay}(i, D_i, g)$. Let $\{C_i\}$ be the collection of all decision procedures which maximize this payoff; the global disambiguation ensures that the C_i's agree on g. $C_i(g)$ is therefore the unique globally rational move in g. □

This is in some sense a surprising result, since it does not require the players to have common knowledge of the payoff function (in that each agent knows that the other agents know that it knows the payoff function, etc.). Halpern and Moses refer to this situation, where everyone knows p, but without the further requirements of full common knowledge, as *E-knowledge*. In light of their result [8] that it is virtually impossible for a distributed system to achieve common knowledge, this is an important consideration.

The following is an immediate consequence of theorem 6.1:

Theorem 6.2. *Suppose that each agent in a globally disambiguated interaction makes the assumption of common behavior, and that the density functions of the various agents match. Then separate global rationality implies that the common behavior assumption will in fact be valid.*

Proof. Theorem 6.1 says that each agent will have at most one globally rational decision procedure under these assumptions; lemma 4.3 guarantees that this globally rational decision procedure is the one which satisfies the common behavior assumption. □

In other words, we do not need to equip interacting agents with decision procedures which are in fact permutations of one another; if we provide them merely with the *assumption* of common behavior and with instructions to behave in a globally rational fashion, their decision procedures will match as a result.

Recall that a move $\vec{m}$ in a game is *Pareto optimal* if, for any move $\vec{n}$, there is a player i for whom $p(i, \vec{n}) < p(i, \vec{m})$. The content of theorem 5.3, for example, is that Pareto-suboptimal moves are irrational under the assumption of common behavior.

Consider now a meta-game M in which a "move" for player i is the selection of a decision procedure D_i, so that a joint move is the selection of a joint decision procedure D_S. We define the payoffs in M by

$$p(i, D_P) = pay(i, D_i);$$

this is the payoff to i of selecting decision procedure D_i (if the other players select the decision procedure D_i). The main result of this section is the following:

Theorem 6.3. *If all players have matching density functions, and all games in G are locally disambiguated, then any collective decision procedure D_P which obeys the common behavior assumption and is globally rational for the players in P will be Pareto optimal in the meta-game M.*

Proof. Again, we can assume without loss of generality that $G = o(g)$ for some fixed game g with payoff function p.

Under the assumption of common behavior, there will be a single move $\vec{c}$ which is made in g, with permutations of $\vec{c}$ being made in permutations of g. Now fix a player i; we will denote by C_i the decision procedure which generates the move $\vec{c}$.

The payoff to i of the decision procedure C_i is now given by

$$\begin{aligned} pay(i, C_i) &= \sum_{g' \in G} \rho_i(g') pay(i, C_i, g') \\ &= \sum_{\sigma \in S_n} \rho_i(\sigma) pay(i, C_i, \sigma(g)) \\ &= \sum_{\sigma} \rho_i(\sigma) p_\sigma(i, C_i(\sigma(g))) \\ &= \sum_{\sigma} \rho_i(\sigma) p_\sigma(i, \sigma(\vec{c})) \\ &= \sum_{\sigma} \rho_i(\sigma) p(\sigma^{-1}(i), \vec{c}). \end{aligned}$$

The local disambiguation assumption guarantees that every move $\vec{c}$ is available as the outcome of a commonly rational decision procedure; C_i will therefore be globally rational just in case

$$\sum_{\sigma} \rho_i(\sigma) p(\sigma^{-1}(i), \vec{c}) \geq \sum_{\sigma} \rho_i(\sigma) p(\sigma^{-1}(i), \vec{d}) \tag{11}$$

for all moves $\vec{d}$ in g.

If C_i were not Pareto optimal, there would be a collection of moves, one for each game in G (i.e., for each $\sigma \in S_n$), leading to better global payoffs for all of the players in g. If we denote these moves by $\vec{m}_\sigma$, the payoff to i will be

$$\sum_{\sigma \in S_n} \rho_i(\sigma) p_\sigma(i, \vec{m}_\sigma).$$

Meanwhile, the payoffs to another player $\sigma'(i)$ will be

$$\sum_{\sigma \in S_n} \rho_{\sigma'(i)}(\sigma) p_\sigma(\sigma'(i), \vec{m}_\sigma).$$

It follows that C_i will not be Pareto optimal only if

$$\sum_{\sigma} \rho_{\sigma'(i)}(\sigma) p_\sigma(\sigma'(i), \vec{m}_\sigma) \geq \sum_{\sigma''} \rho_i(\sigma'') p(\sigma''^{-1}(i), \vec{c})$$

for all σ', with the inequality being strict at least once. (The global payoff under common behavior does not change from player to player). This implies

$$\sum_{\sigma'}\sum_{\sigma}\rho_{\sigma'(i)}(\sigma)p_\sigma(\sigma'(i),\vec{m}_\sigma) > n!\sum_{\sigma''}\rho(\sigma'')p(\sigma''^{-1}(i),\vec{c}). \tag{12}$$

The left hand side of (12) can be easily rewritten as

$$\sum_{\sigma}\sum_{\sigma'}\rho_{\sigma'(i)}(\sigma)p_\sigma(\sigma'(i),m_\sigma) = \sum_{\sigma}\sum_{\sigma'}\rho_i(\sigma'^{-1}\sigma)p(\sigma^{-1}\sigma'(i),\sigma^{-1}(m_\sigma)).$$

If we write σ'' for $\sigma'^{-1}\sigma$ and m'_σ for $\sigma^{-1}(m_\sigma)$, this is now

$$\begin{aligned}\sum_{\sigma}\sum_{\sigma'}\rho_{\sigma'(i)}(\sigma)p_\sigma(\sigma'(i),m_\sigma) &= \sum_{\sigma}\sum_{\sigma''}\rho_i(\sigma'')p(\sigma''^{-1}(i),m'_\sigma)\\ &\leq \sum_{\sigma}\sum_{\sigma''}\rho_i(\sigma'')p(\sigma''^{-1}(i),\vec{c})\\ &= n!\sum_{\sigma''}\rho_i(\sigma'')p(\sigma''^{-1}(i),\vec{c})\end{aligned}$$

This is in conflict with (12), and the proof is complete. □

I cannot resist the temptation to rephrase this: if all men are indeed created equal, the Golden Rule is valid. Well, at least it's Pareto optimal.

§7. Communication

We seem to have drifted rather far from the concerns of DAI. This is not the case, though; the point of this paper is the following:

Proposal. *Intelligent agents should be equipped with the common behavior assumption.*

There are a variety of reasons for this. The first is the Pareto optimality of common behavior, as in theorem 6.3. Provided that the agents being considered expect to encounter a similar distribution of interactive situations, there is *no* assumption or set of assumptions which would make the agents uniformly more effective problem solvers.

It can be argued that the assumption of matching density functions is an invalid one, and this is a fair criticism. In the presence of specific information regarding differences in the density functions of various agents, it is possible that more effective assumptions can be found. But in many cases, the strength of the local results in section 5 may be sufficient to ensure a satisfactory result if the common rationality assumption is adopted.

Additional advantages are a consequence of theorem 6.1. *Provided that interacting agents can agree on a globally disambiguated payoff function for a game g, they need communicate no further.* This is a consequence of the fact that this payoff function, in combination with the common behavior assumption, will be sufficient to determine the agents' actions.

There are a variety of advantages to restricting communication in this fashion. Firstly, it is possible to establish uniform protocols for the exchange of information regarding payoff matrices. One such protocol would simply have the agents broadcast their version of the payoff function in sequence; as soon as there is uniform agreement on a value, the interaction could proceed.

The reason that we are able to get by without *common* knowledge of the payoff function will be a bit clearer if we return to our widget example. Let us suppose that the Palo Alto robot broadcasts a payoff matrix indicating the advisability of stopping to pick up the various tools, and that the Boston robot confirms this. At this point, the robots know the following:

(1) Both robots know that it is to their advantage to stop and acquire the tools, and
(2) The Palo Alto robot knows that the Boston robot knows that this is the case.

It is *not* the case, though, that the Boston robot knows that the Palo Alto robot is aware of the advantages to be gained by acquiring the tools, since the Palo Alto robot has not yet acknowledged the Boston robot's message confirming the payoff function (nor will he); all that the Boston robot knows is that *when the Palo Alto robot receives the Boston robot's last message*, he will be aware of the values in the confirmed payoff matrix.

The effect of this is that the Boston robot does not know precisely *when* the Palo Alto robot will be prepared to start on his journey to Pittsburgh; all he knows is that when he arrives, he will have a borogove with him.

Let us return to more general considerations. In the presence of disagreement, for example, it will often be possible for one agent to *prove* its description of the payoff function to be correct. Such a proof might simply explain the situation in which that particular agent finds itself, or might instead provide a useful medium for agents to advise one another about possibly unforeseen consequences of their activities. If the Boston robot in our widget example had been unaware of the value of having a zazzyfrax and a borogove to help with the construction, the demonstration by the Palo Alto robot that there was mutual benefit to be had by stopping and picking up these tools would constitute advice of just this sort.

The fact that it is possible to demonstrate the accuracy of a payoff function will also alleviate problems arising from the fact that it may be to a particular agent's advantage to lie, as has been remarked by Rosenschein [15]. It is awkward to prevaricate in a situation where one can be required to vindicate one's claims! The only other solution which has been proposed to this, again by Rosenschein [14], is to have agents exchange binding promises, but it is not at all clear how the binding nature of these promises could be effected.

There is one point we have left unaddressed, and that involves the practical import of the disambiguation assumptions in the theorems of the last section. Fortunately, the following result should be fairly clear:

Theorem 7.1. *Every game is arbitrarily close to one which is locally and globally disambiguated.* □

Note, however, that since in some cases the decision procedure produced by theorem 6.1 may vary discontinuously with respect to the payoff function for the game in question, it is possible for a group of intelligent agents to have difficulty agreeing on a perturbation

that would disambiguate it. It does not appear that such difficulties can be addressed within the framework we have developed.

§8. An application

Let me end with an example that will hopefully make the practical import of these ideas a bit clearer. I will imagine a very simple environment which contains two processors and one peripheral which can be used by the processors to speed their computation. One of the processors is initially responsible for dividing a problem into two smaller ones; it then hands one of these subgoals off to the other machine. The peripheral might be, for example, a floating point accelerator (FPA).

The two processors (which I will call simply 1 and 2) must now decide whether or not to request access to the FPA; each processor has the option of making a request (R) or of making no request (N). In the interest of simplicity, I will assume that there is no facility for arbiting between conflicting requests: if both processors request the peripheral, both are denied access.

The payoffs to the two processors are related inversely to the amount of time that they will take to complete their respective tasks. I will assume, though, that the payoff to each processor is simply the amount of time taken for its computation, with the goal of the processors now being to *minimize* their payoffs.

Initially, then, each processor receives a task, and will determine what the advantage is of requesting access to the FPA. Let me assume that each processor makes this calculation for itself and transmits the results. If processor 1 (the "row" player) had payoffs of 6 (without the FPA) and 2 (with it), while processor 2 had payoffs of 4 and 1, the payoff matrix would look like this:

	R	N
R	4 6	4 2
N	1 6	4 6

(13)

Note first that this game is a version of the prisoner's dilemma. Since each processor is no worse off requesting access to the FPA than not, an application of case analysis will lead to the unfortunate choice $\{RR\}$.

Let us suppose for the moment that the density functions ρ_i are uniform over all games. An examination of (11) now reveals that any commonly rational move will be one which maximizes the total payoff to all of the players involved. This leads to the joint move RN above; the processor which will benefit the most from the use of the FPA requests it. The effect of this is to minimize the total amount of time needed for the two processors to complete their computations (i.e., they use the fewest number of cycles).

Let us now allow the two processors to do a little more meta-level reasoning. They realize that if processor 1 does not request the FPA, but waits until the other processor is done with it before starting its own computation, it will actually complete its task in 3 units of time, as opposed to 6. The second processor can reason similarly, and the payoff matrix should therefore be replaced with this one:

	R	N
R	6, 4	2, 3
N	3, 1	6, 4

(14)

Common rationality now leads to the joint move NR; the second processor requests the FPA while the first waits for him to finish using it.

Other density functions can be used to produce different results. If, for example, it is desired to have the slowest processor complete its task in the shortest time possible, the density functions need merely be taken to be very sharply increasing functions of the time needed to complete the task without access to the FPA. If it is desired to have the *fastest* processor complete its task as quickly as possible (perhaps to begin consideration of another problem), the density functions should be very sharply *decreasing* functions of the time needed to complete the task *with* access to the FPA. An examination of (11) will make both of these claims clear. Of course, all of these methods can be used just as easily with the revised payoff function (14) as with the earlier version (13).

Let me conclude by describing the advantages of the common rationality approach over a few other possibilities. The one that springs to mind most quickly is that of having the task-assigning machine determine which of the subprocessors should use the FPA. This will be reasonable in simple situations, but there may well be a substantial amount of analysis involved in arriving at the payoff functions appearing in (13), and this analysis can be usefully distributed between the other machines. It should not be argued, however, that this substantial analysis is a "waste of time"—Smith has pointed out [16] that meta-analysis will become increasingly important as machines tackle more and more difficult problems.

The attractiveness of common rationality is also not merely a consequence of the harsh penalty we have imposed if both processors request access to the FPA. Suppose that the FPA is merely assigned at random in such a case; (13) should then be replaced with:

	R	N
R	4, $2\frac{1}{2}$	2, 4
N	6, 1	6, 4

Individual rationality still forces the move $\{RR\}$; although this move is no longer Pareto-suboptimal (as it was earlier), there may nevertheless be many situations in which we want to avoid it.

A final possibility would be to have the processors agree on a payoff function and to then have one of them determine what action should be taken. The point of theorem 6.2, however, is that this is unnecessary: the various processors will coordinate their actions automatically, without *needing to select a single machine* to arbitrate their decisions.

Not only is this additional complication unnecessary, but there are many situations where it will be outright impossible—the interaction may be unexpected, or may be between agents with conflicting goals. There are a variety of examples in [5], although the state of research into the problems of interacting intelligent agents seems to preclude describing them quantitatively at this point.

§9. Conclusion

The decision procedure paradigm is a new one in which to examine multi-agent interactions: the results of this paper should be novel to both game-theoretic and AI audiences.

The advantage of this paradigm is that it makes it possible to avoid the usual assumption that the decision methods of the various agents are independent. This independence can be assumed, but it is only one of a variety of assumptions which can be made.

One of these, which we have referred to as common rationality, seems especially attractive. It enables us to understand coordinated action in terms of local utility maximization, and solves the prisoner's dilemma. We have also showed common rationality to be Pareto optimal among the collection of all possible assumptions which might be made by interacting agents about each other.

The common rationality assumption is also strong enough to limit the communication needed between such interacting agents. By allowing the agents to communicate information describing their situation instead of describing their proposed intentions, we can establish a uniform protocol for this communication while avoiding the difficulties that would otherwise be a consequence of potential agent dishonesty.

The difficulties with the common rationality approach appear to lie in two assumptions: that the interacting agents have matching density functions, and that any particular game not be locally or globally ambiguous. Only the first of these is likely to present a serious difficulty in practice, and future research will need to address interaction between agents with varying expectations about the situations in which they will find themselves.

Acknowledgement

The fact that I have occasionally chosen to disagree with the conclusions that Mike Genesereth and Jeff Rosenschein have drawn elsewhere is unrepresentative of the influence they have had on this work. Discussions with them have been invaluable to my development of any thoughts in this area; many of the ideas in this paper appeared initially in a paper authored jointly with them [5]. Joe Halpern helped by nagging me incessantly until all of the gritty details involving the permutations had been sorted out.

This research has been supported by the Office of Naval Research under grant number N00014-81-K-0004.

References

[1] Axelrod, R., *The Evolution of Cooperation*, Basic Books, Inc., New York (1984).

[2] Corkill, D.D., and Lesser, V.R., The use of meta-level control for coordination in a distributed problem solving network, *IJCAI-83*, Karlsruhe, West Germany (1983) 748–756.

[3] Davis, R., A model for planning in a multi-agent environment: steps toward principles for teamwork, Working Paper 217, MIT AI Lab (1981).

[4] Davis, R., and Smith, R. G., Negotiation as a metaphor for distributed problem solving, *Artificial Intelligence* 20 (1983) 63–109.

[5] Genesereth, M.R., Ginsberg, M.L., and Rosenschein, J.S., Cooperation without communication, HPP report 84-36, Stanford University (1984).

[6] Georgeff, M., Communication and interaction in multi-agent planning, *AAAI-83*, Washington, D.C. (1983) 125–129.

[7] Georgeff, M., A Theory of action for multi-agent planning, *AAAI-84*, Austin, Texas (1984) 121–125.

[8] Halpern, J., and Moses, Y., Knowledge and common knowledge in a distributed environment, *Proceedings of the Third Annual ACM Conference on Principles of Distributed Computing*, Vancouver, British Columbia, Canada (1984).

[9] Hofstadter, D.R., Metamagical themas—the calculus of cooperation is tested through a lottery, *Scientific American* 248 (1983) 14–28.

[10] Konolige, K., A first-order formalization of knowledge and action for a multi-agent planning system, Tech Note 232, SRI International, Menlo Park, California (1980).

[11] Luce, R.D., and Raiffa, H., *Games and Decisions, Introduction and Critical Survey*, John Wiley and Sons, New York (1957).

[12] Malone, T W., Fikes, R.E., and Howard, M.T., Enterprise: a market-like task scheduler for distributed computing environments, Working paper, Cognitive and Instructional Sciences Group, Xerox Palo Alto Research Center (1983).

[13] Rapoport, A. and Guyer, M., A taxonomy of 2 x 2 games, *Yearbook of the Society for General Systems Research* XI (1966) 203–214.

[14] Rosenschein, J.S. and Genesereth, M.R., Deals among rational agents, IJCAI-85, Los Angeles, California (1985).

[15] Rosenschein, J.S., *Rational Interaction: Cooperation Among Intelligent Agents*, Ph.D. thesis, Stanford University (1985).

[16] Smith, D.E., *Controlling Inference*, Ph.D. thesis, Stanford University (1985).

END

FILMED

2-86

DTIC